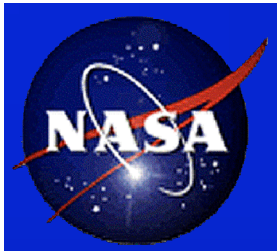


Single-Trial Characterization of Evoked Responses

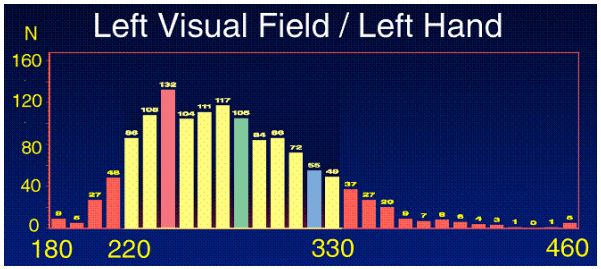
Bayesian Estimation and Differentially Variable Component Analysis

Kevin H. Knuth
NASA Ames Research Center
Moffett Field CA, USA

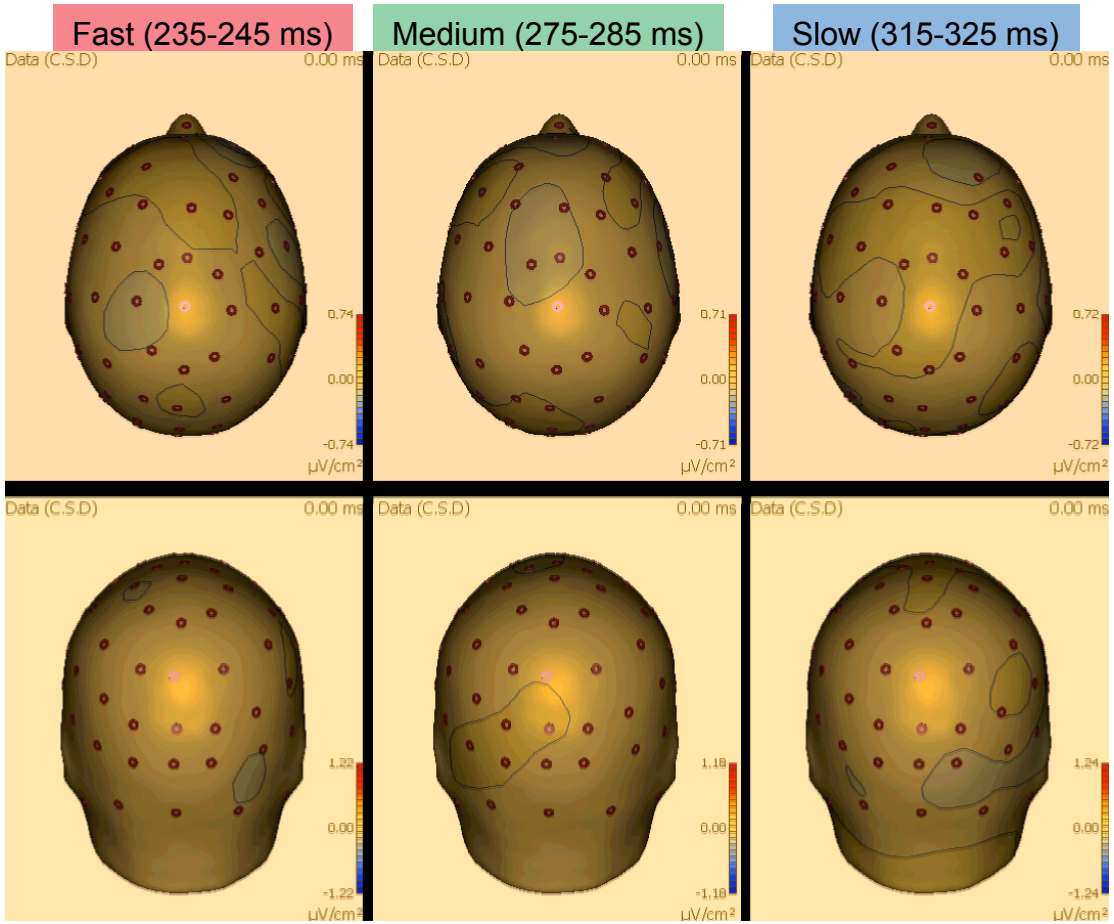
Brains are Dynamic

Brains are dynamic, state dependent, and change over time.

Reaction Time Histogram



Reaction Time



Movie courtesy of Clifford Saron.

Saron C.D. et al. 2003. In: *The Asymmetrical Brain.* Hugdahl, K and Davidson, R.J., Eds. Pp. 341-408.

Intracranial ERP Recordings

Intracranial Recordings

Multi-Electrode Array

Span Cortical Layers

In these data:

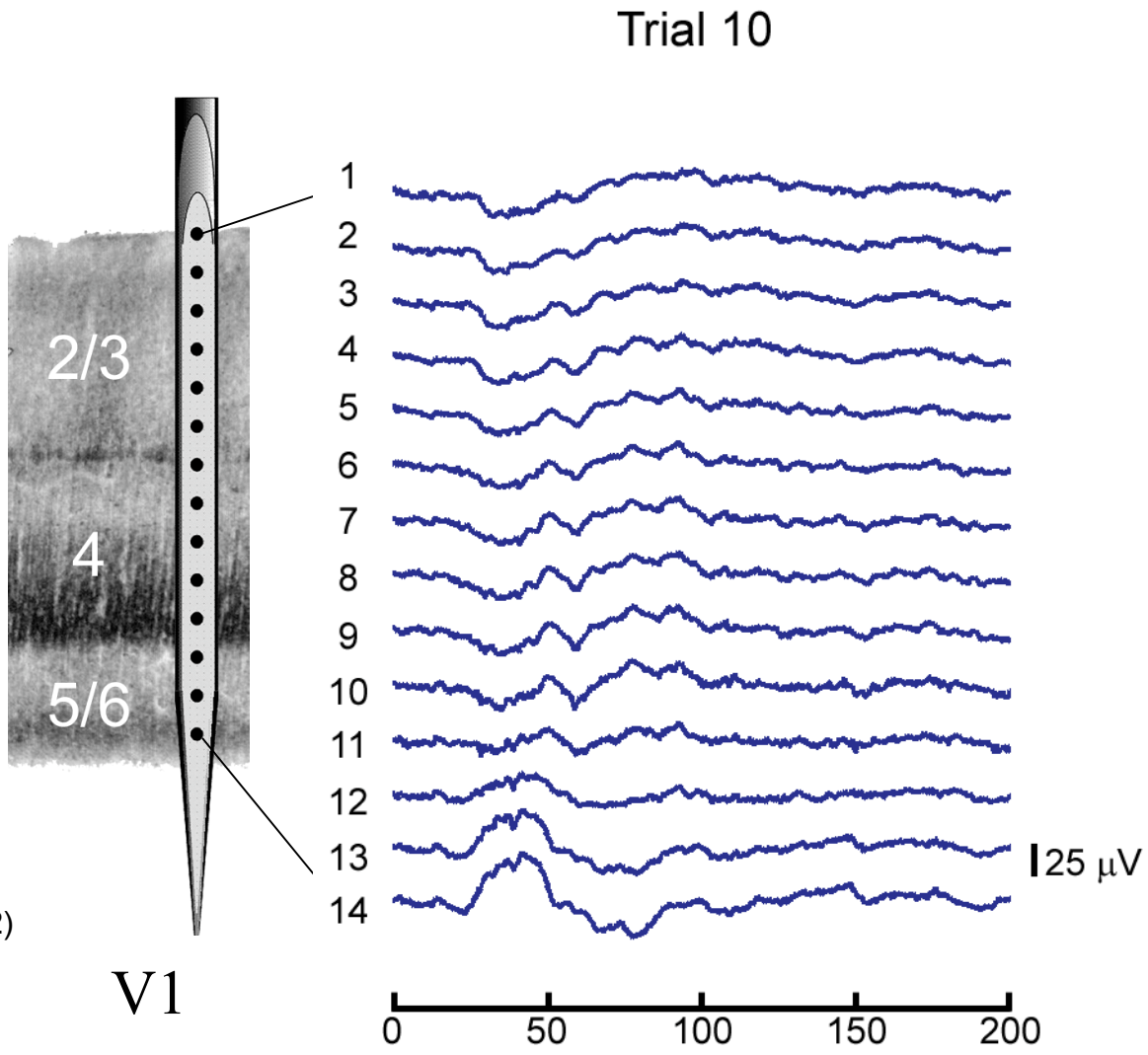
1. Evoked responses

2. No significant phase resetting

Shah AS, et al. 2004. In press: *Cereb Cortex*

3. Trial-to-trial variability

Truccolo WA, et al. 2002. *Clin Neurophysiol* 113(2)

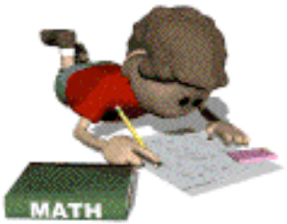


Averaging ERPs

Under what conditions is averaging single trial data appropriate?

If we assume that the ERP is identical from trial-to-trial, and the data recorded from each trial consists of a sum of this ERP and ongoing background activity and noise, then our signal model for data recorded during trial r is

$$x_r(t) = s(t) + \eta_r(t)$$



Recorded Signal during the r^{th} trial

ERP waveform

Unpredictable Signal Component during the r^{th} trial (Ongoing Activity plus Noise)

Maximum Likelihood Estimation

By assigning a Gaussian likelihood, we can find the most probable solution by maximizing

$$p\left(x(t) \mid s(t), I\right) = \left(2\pi\sigma^2\right)^{-\frac{RT}{2}} Exp\left[-\frac{1}{2\sigma^2}\sum_{r=1}^{R}\sum_{t=1}^{T}\left(x_r(t) - s(t)\right)^2\right]$$

Looking at the Log Probability

$$Log\ p = -\frac{1}{2\sigma^2}\sum_{r=1}^{R}\sum_{t=1}^{T}\left(x_r(t) - s(t)\right)^2 + const$$

We take the derivative with respect to the ERP waveform at a time point q

$$\frac{\partial\ Log\ p}{\partial\ s(q)} = \sigma^{-2}\sum_{r=1}^{R}\left(x_r(q) - s(q)\right) \equiv 0$$

Which setting to zero and solving for the ERP waveform gives $\hat{s}(q) = \frac{1}{R}\sum_{r=1}^{R} x_r(q)$

Thus averaging the trials is optimal!

Response Variability

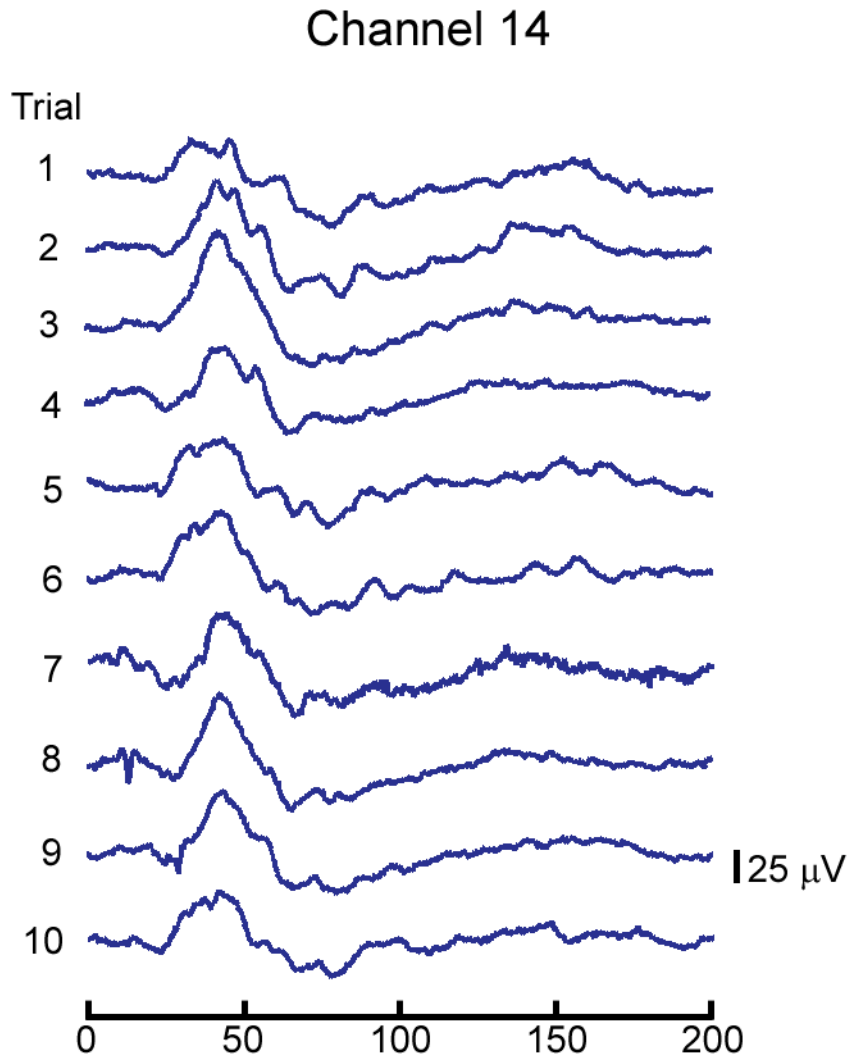
However, the responses are not identical from trial to trial!

There is significant variability:

- Amplitude variations
- Latency variations
- Waveshape variations

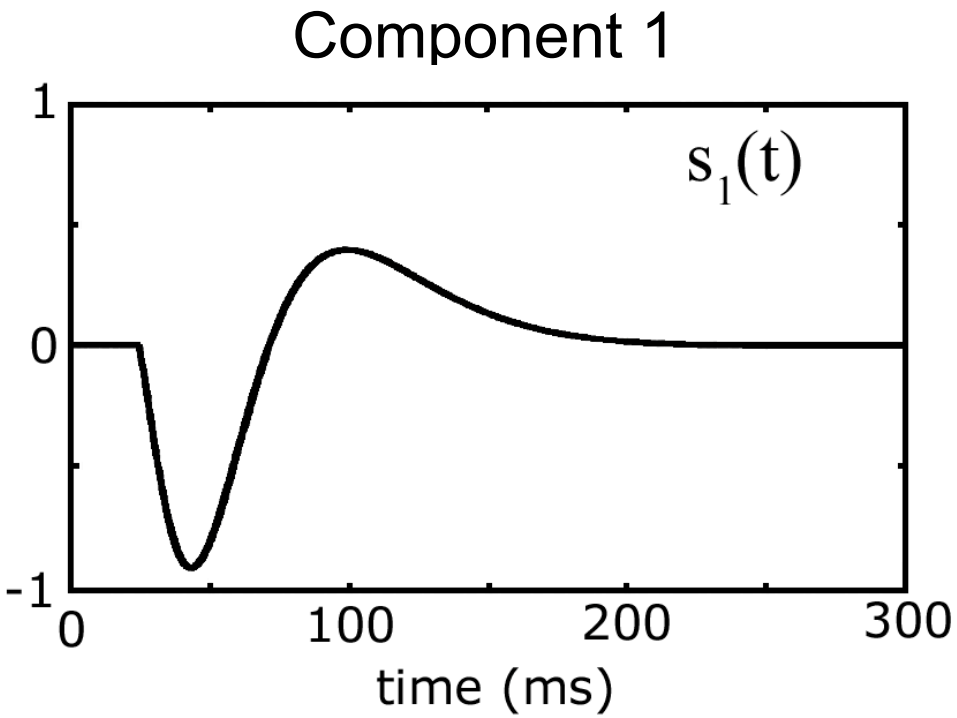
There are also multiple contributing components.

These variations are INTERESTING.



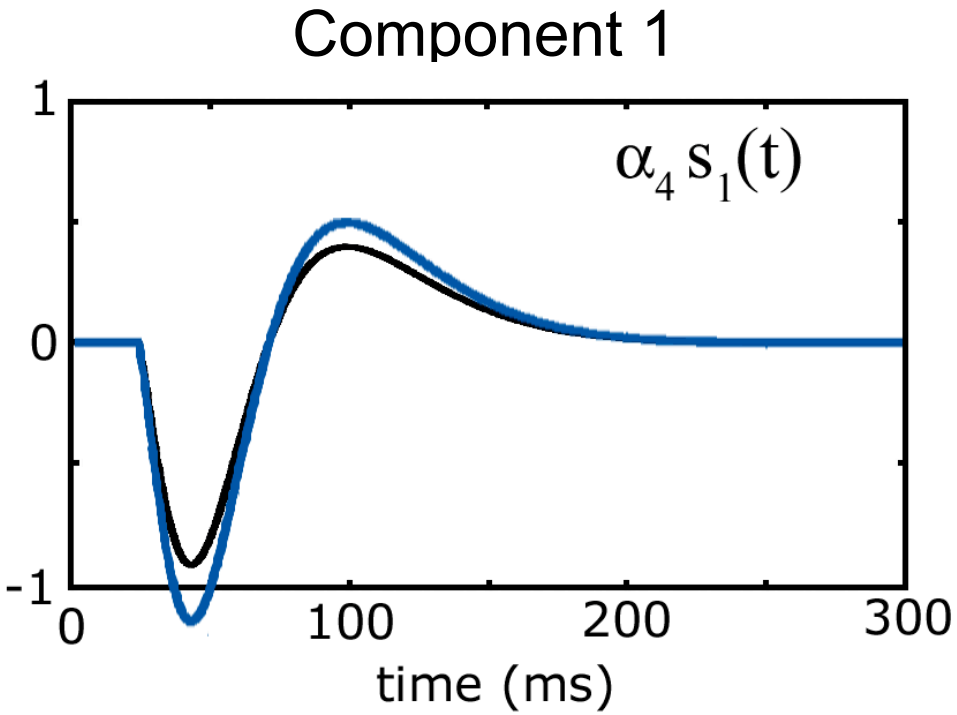
Describing Trial-to-Trial Variability

We begin by describing the waveshape of a single component as a function of time.



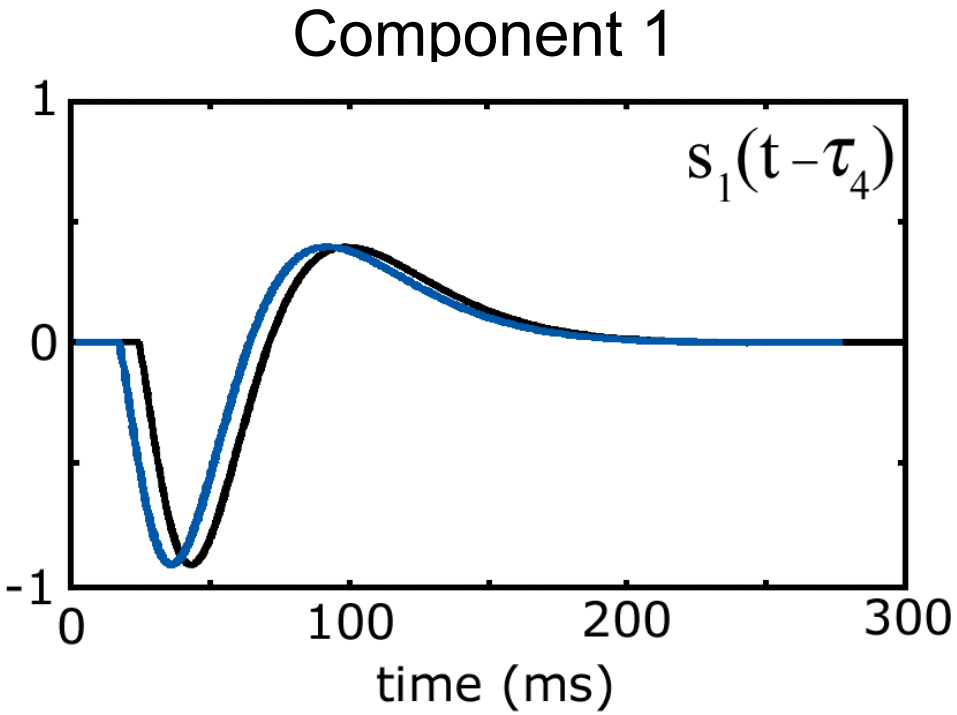
Describing Amplitude Variability

Amplitude can vary from trial-to-trial



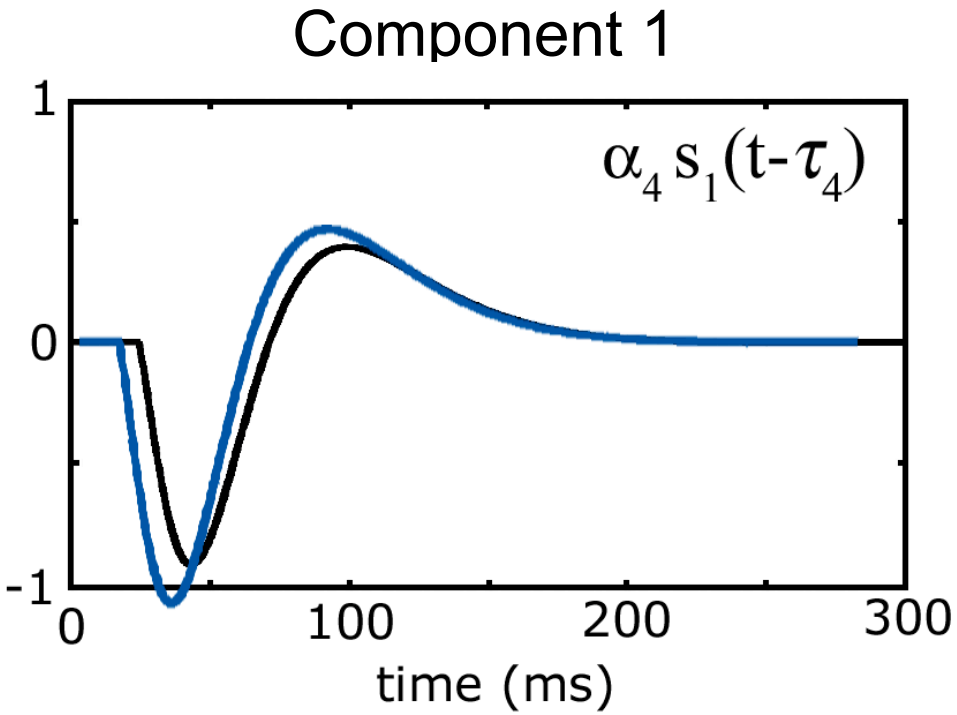
Describing Latency Variability

Onset latency can vary from trial-to-trial



Amplitude and Latency Variability

Now we describe amplitude and latency variability



Multiple Component ERP Model

In this model we assume that each ERP component has its own stereotypic waveshape that can vary in both amplitude and latency from trial to trial. In addition, we allow for multiple channel recordings.

$$x_{mr}(t) = \sum_{n=1}^{N} C_{mn}\,\alpha_{nr}\,s_n\left(t-\tau_{nr}\right) + \eta_{mr}(t)$$

Recorded Signal in the m^{th} channel during the r^{th} trial

Coupling between the n^{th} source and the m^{th} channel

Amplitude of the n^{th} component during the r^{th} trial

$\langle \alpha_n \rangle = 1$

Stereotypic waveform of the n^{th} component

Latency of the n^{th} component during the r^{th} trial

$\langle \tau_n \rangle = 0$

Learning the Parameters from Data

The goal is to use the recorded data to learn the model parameters.

We do this by designing a machine learning algorithm that uses Bayesian probability theory to find the most probable values of the model parameters, such as $s(t)$, α, and τ, given the recorded data.

I have already shown how the most simple model leads to averaging.

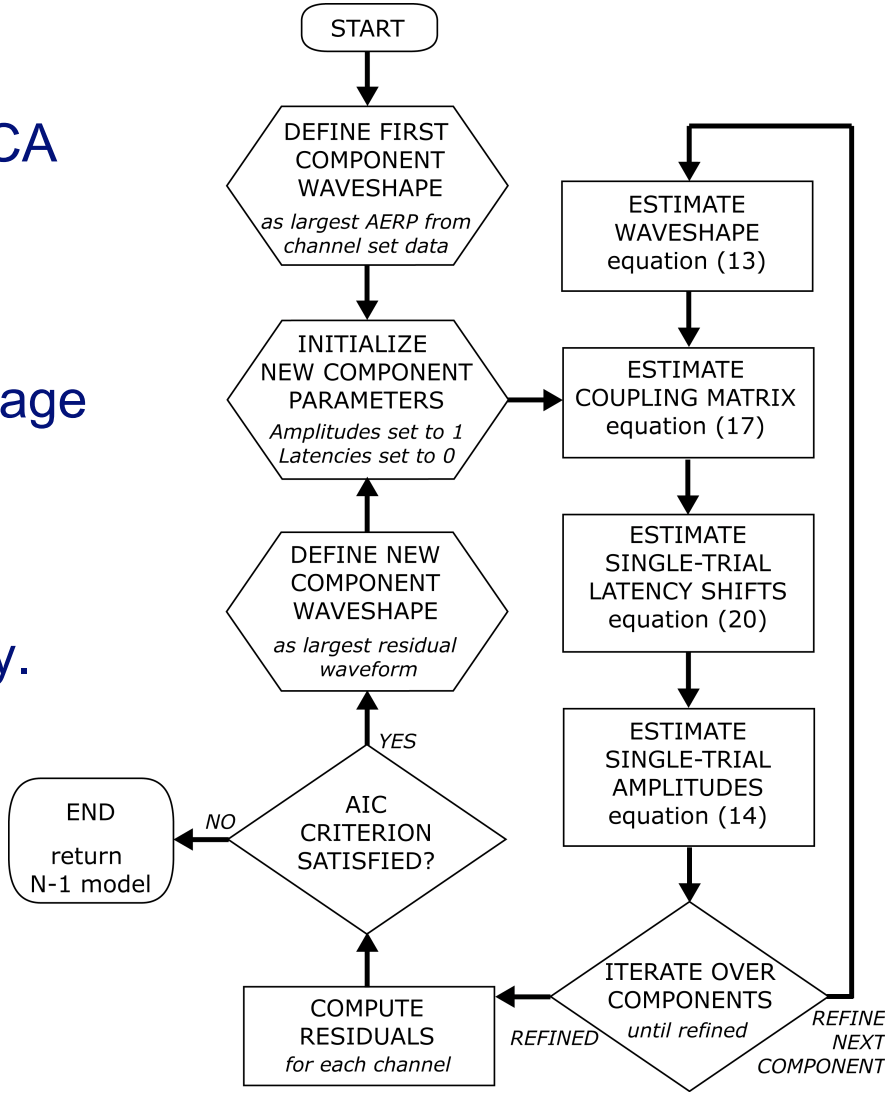
A mixing model that does not take into account variability, but assumes an amplitude density for the components leads to Independent Component Analysis (ICA). Knuth K.H. 1999. *ICA'99.*

Differentially Variable Component Analysis

This is only one of a number of possible ways to implement the dVCA algorithm.

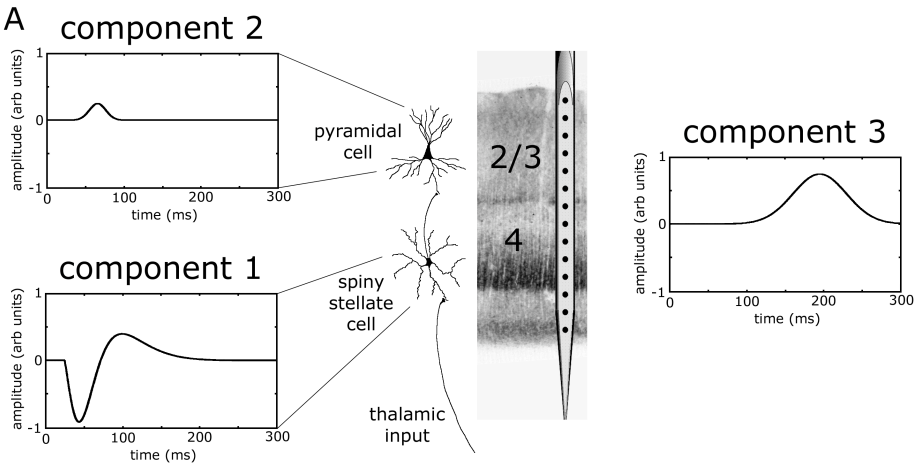
This technique begins with the average ERP and improves it.

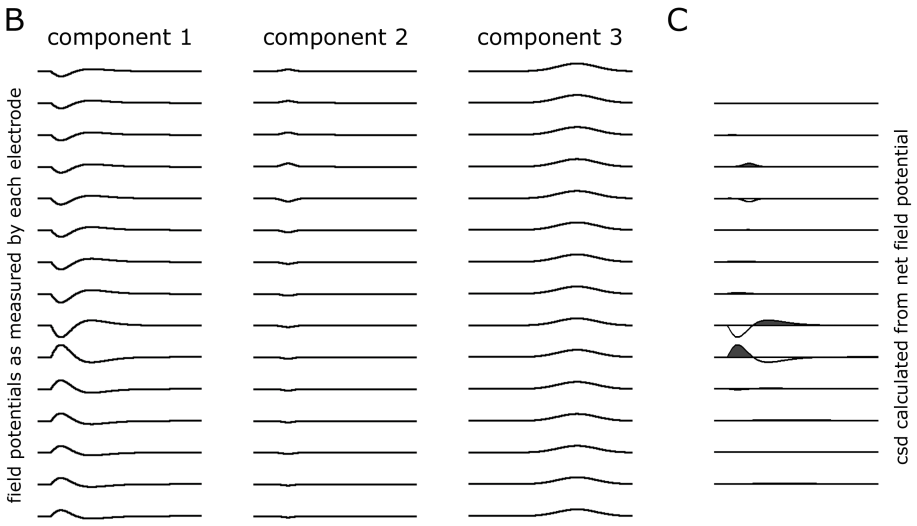
Components are added sequentially.



Simulations

Synthetic Data





To characterize dVCA we created a synthetic dataset designed after the sources hypothesized from our earlier results in Macaque V1 in response to a Red Light Flash.

c1 Initial Layer IV response

c2 Granular response

c3 Far-field response

A Anatomical correlates

B Recorded waveshapes

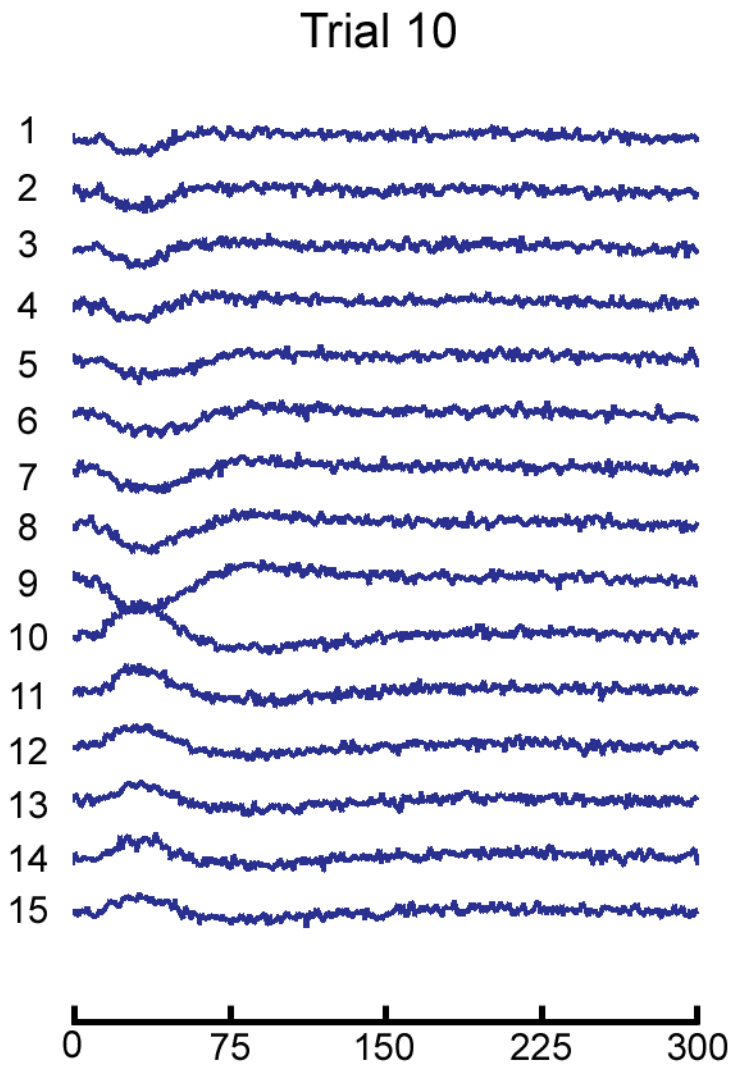
C Current Source Density (CSD)

Multiple Synthetic Trials

Synthetic Trials

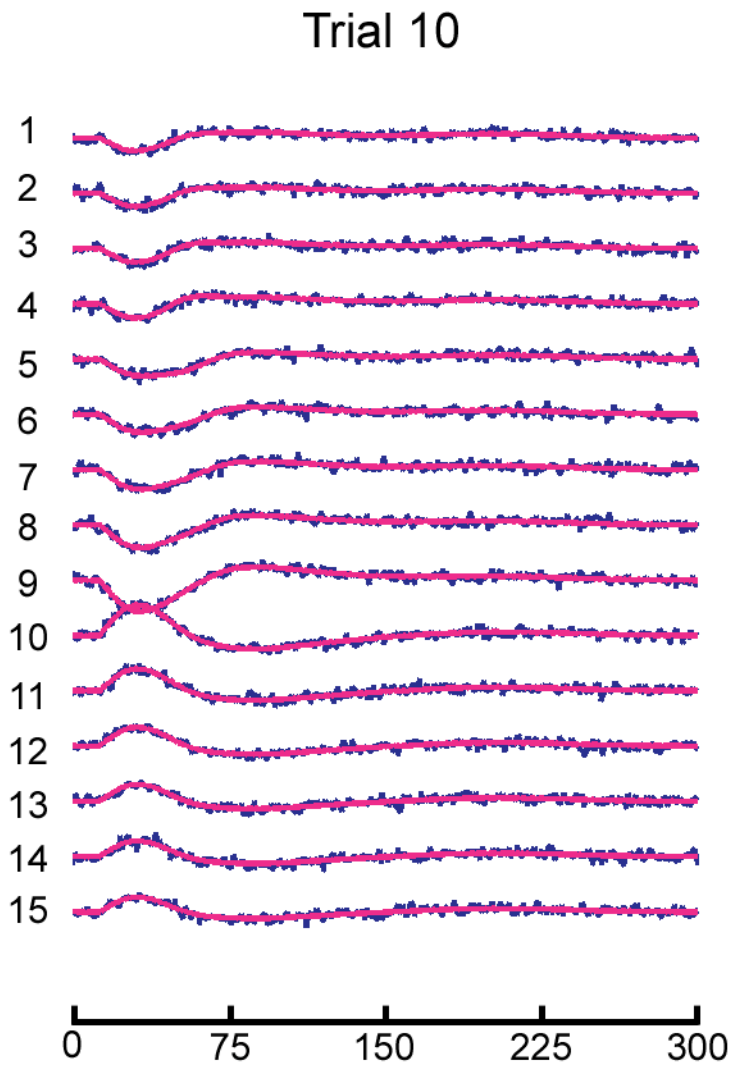
In this example the SNR is low.

All other simulations had higher SNR.

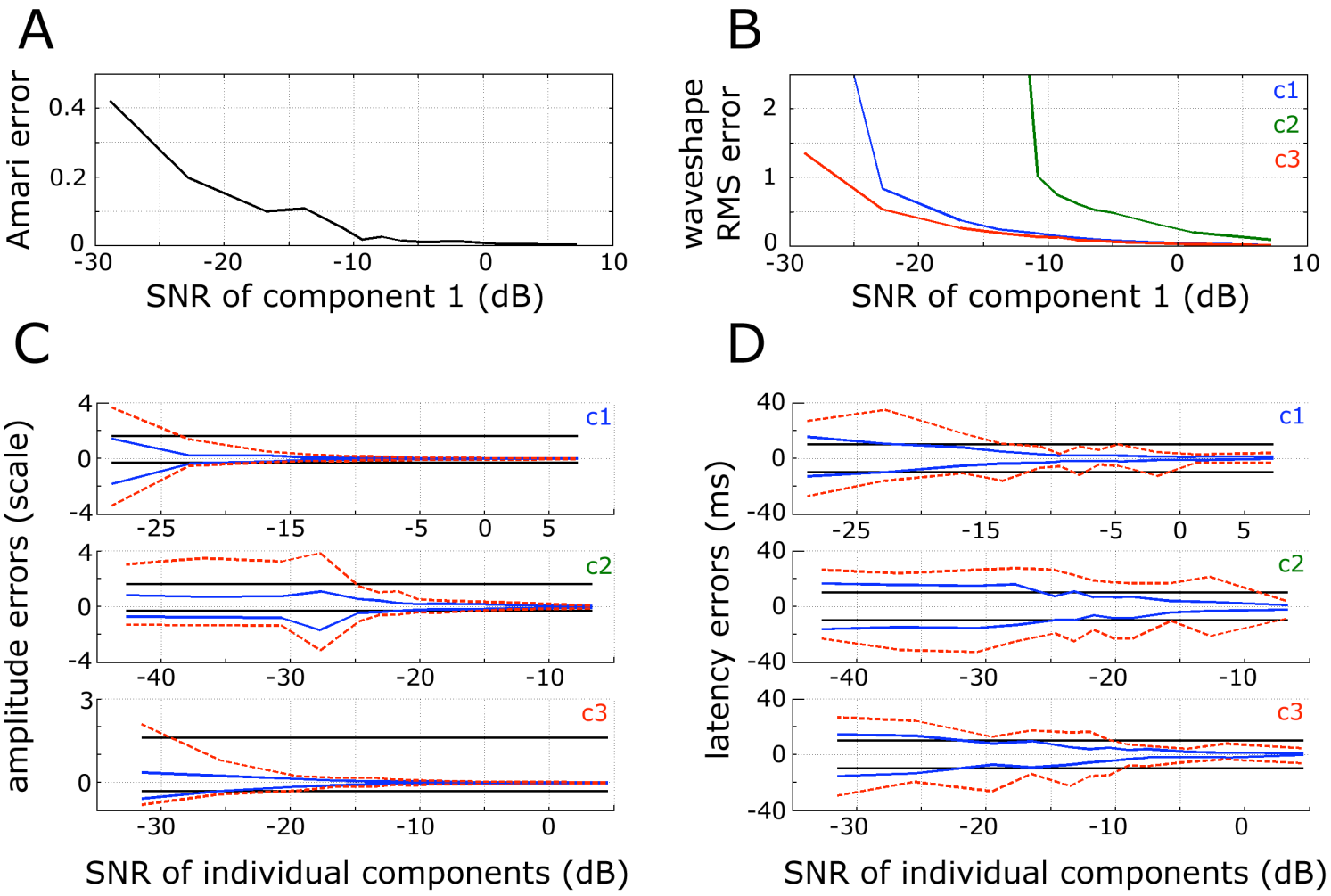


Single-Trial Reconstructions

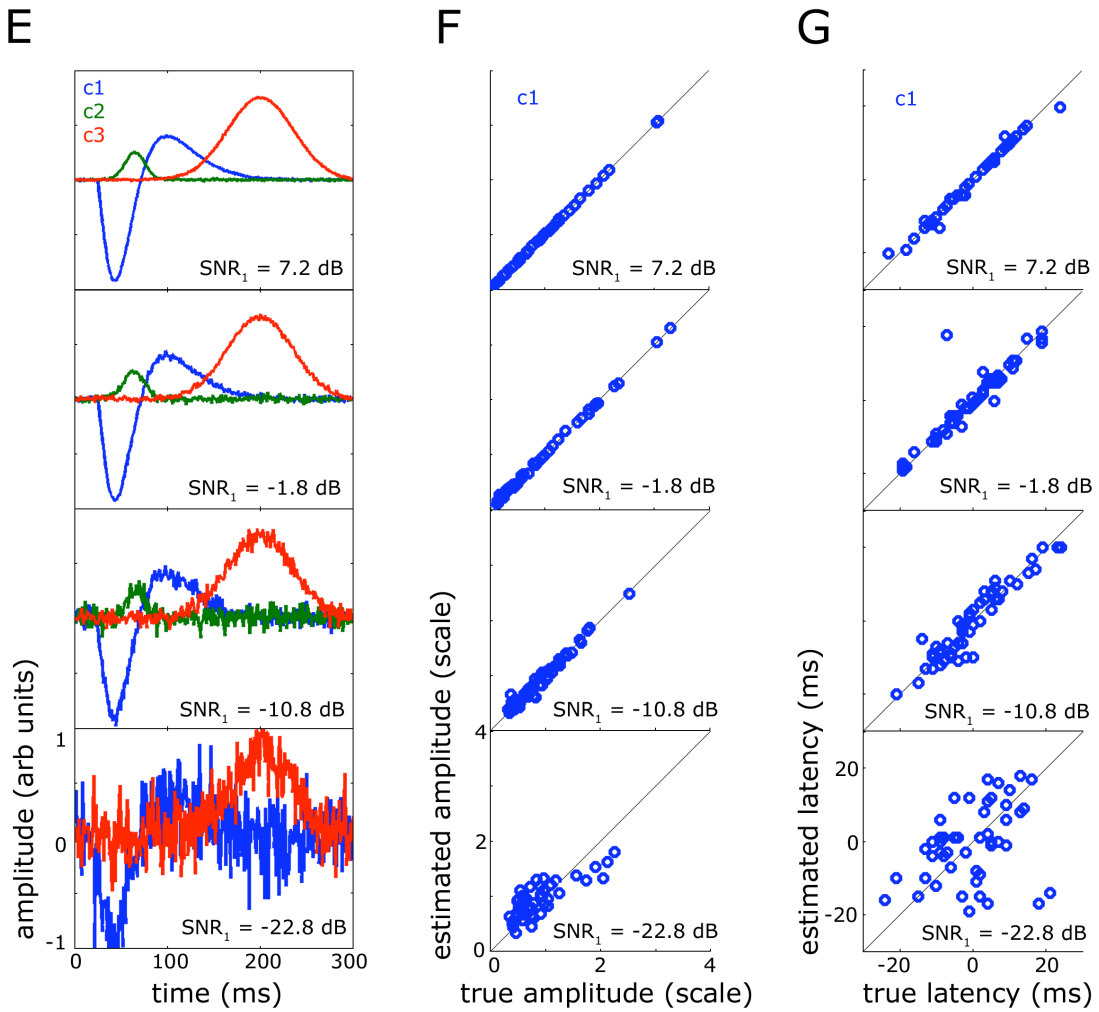
These are the synthetic trials with their reconstructions.



Robustness to Noise



Robustness to Noise



Signal : Noise

5 : 1 (5.28)

2 : 3 (0.67)

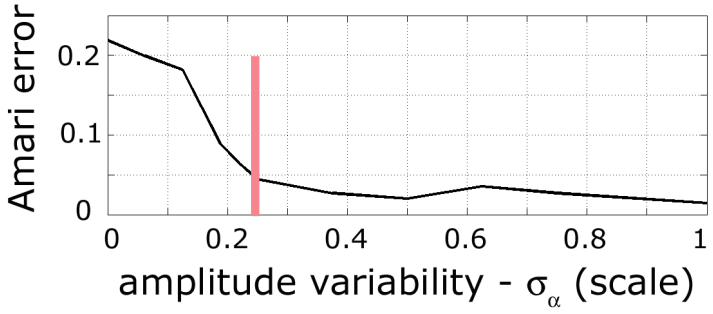
1 : 12 (0.084)

1 : 189 (0.0053)

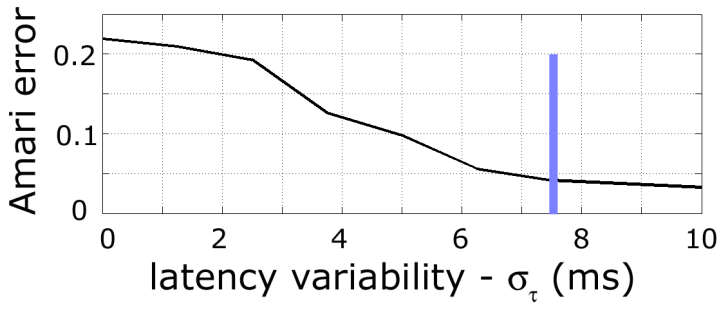
Dependence on Variability

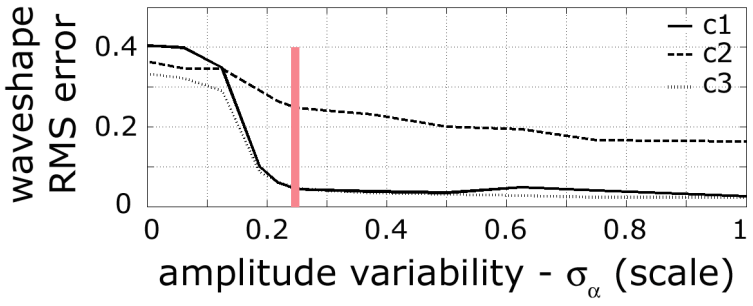
Estimation error decreases with increasing variability

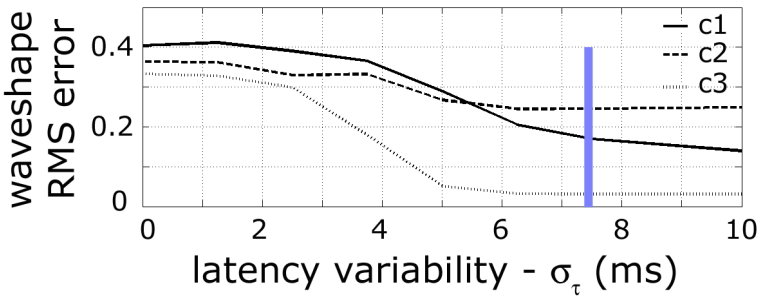
Amplitude Variability



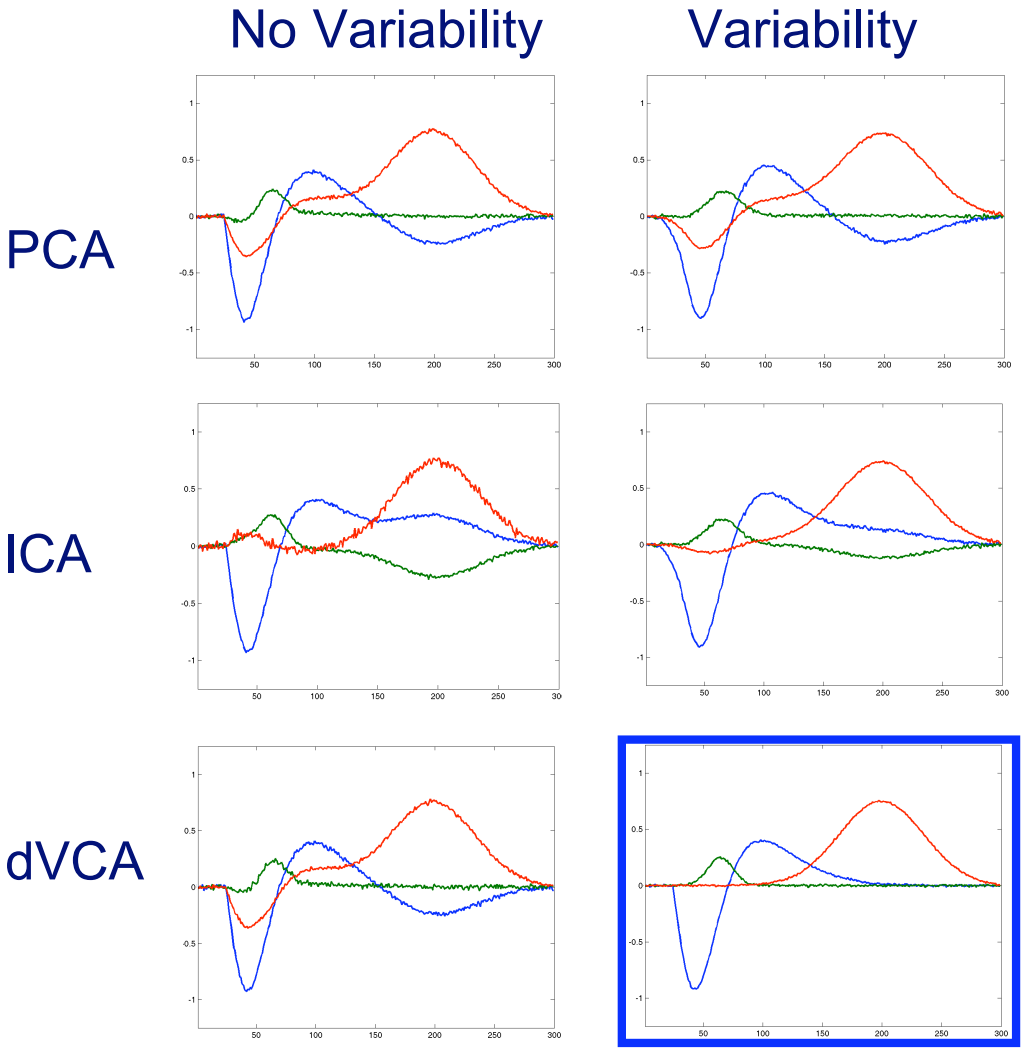
Latency Variability

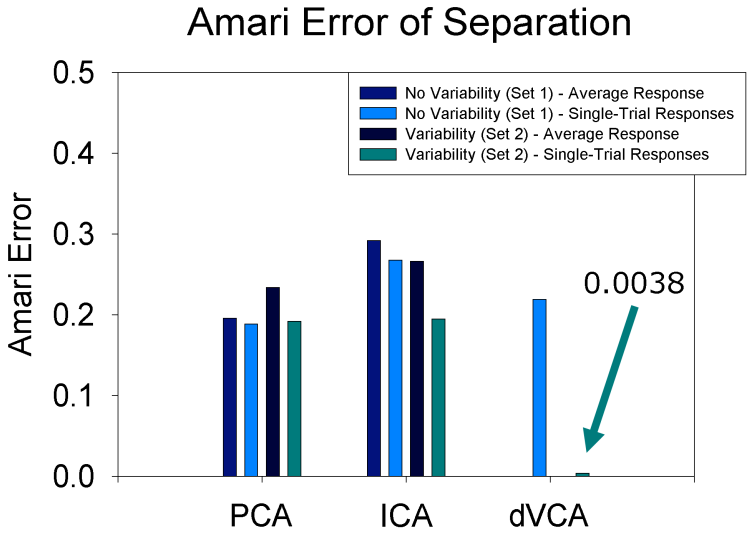


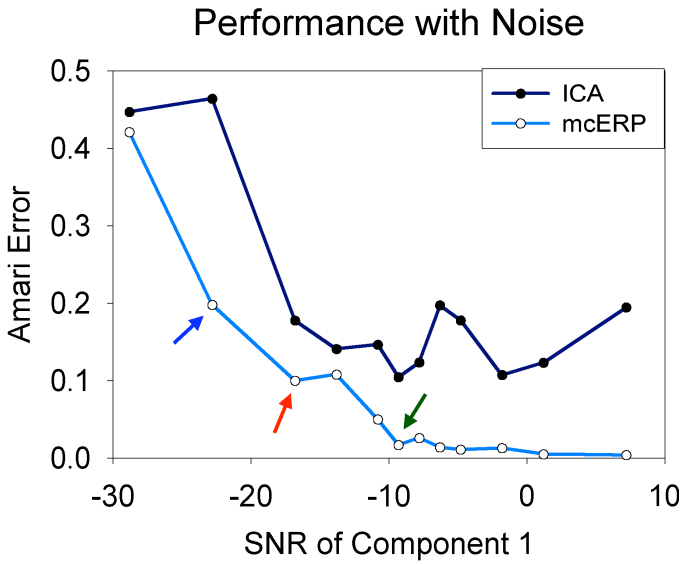




Relation to PCA and ICA







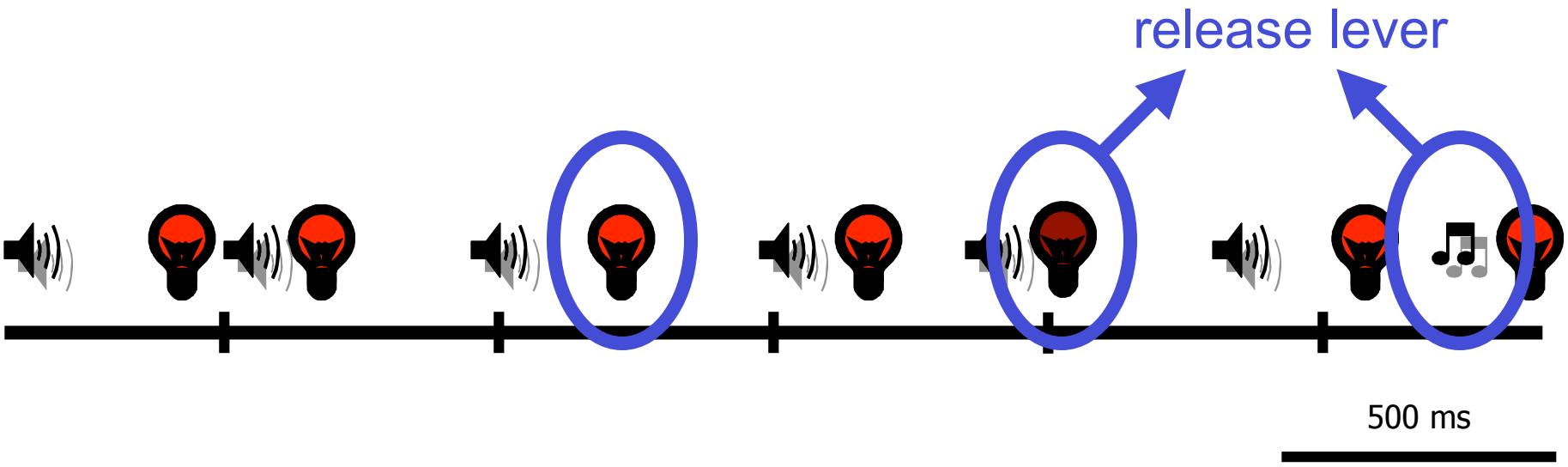
Experimental Data

Application to Real Data

Subject: Male macaca fascicularis

Task: Intermodal selective attention task

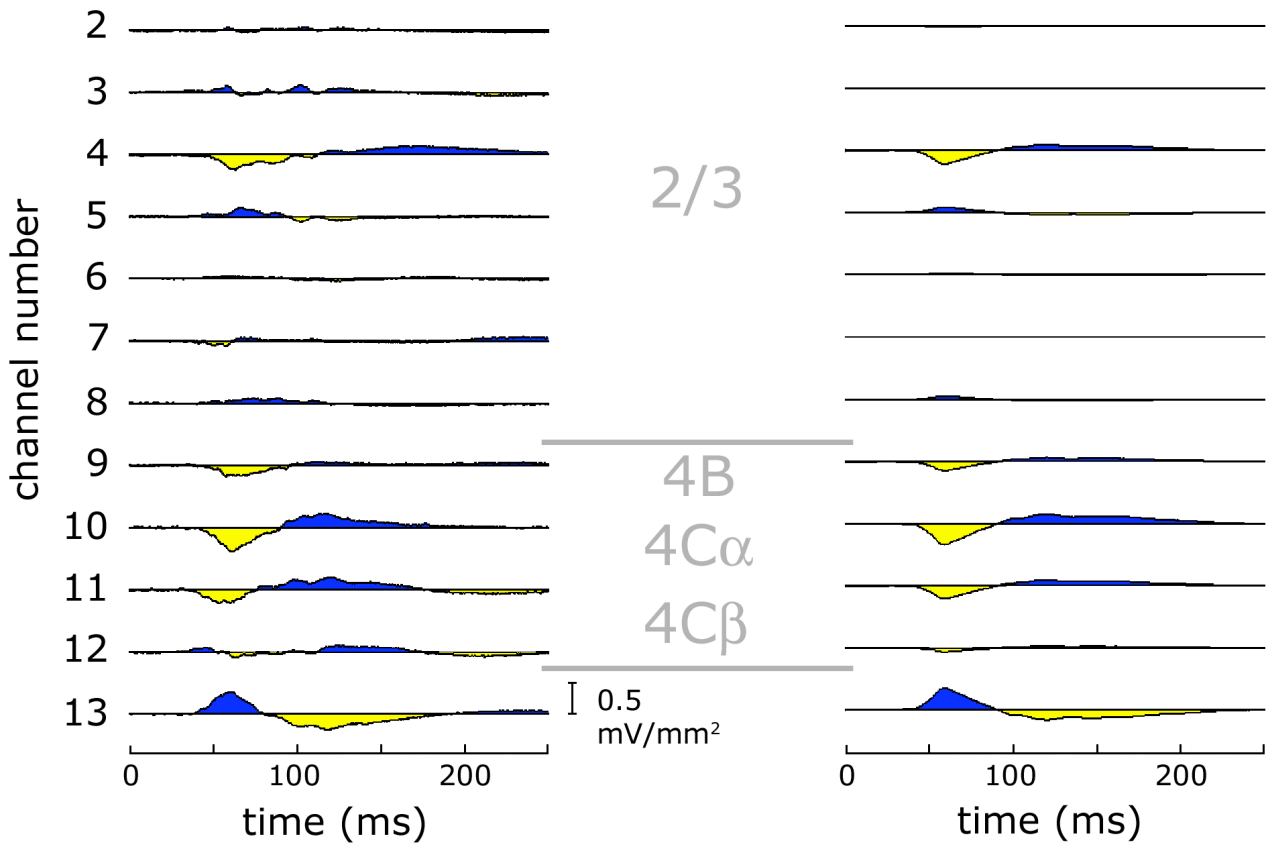
Mehta AD, et al. 2000 Cereb Cortex 10(4)



Estimating a Single Component

CSD of the Average ERP

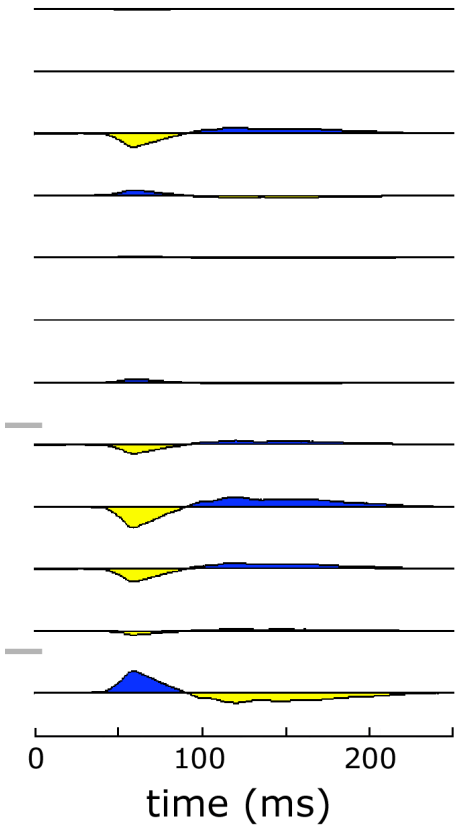
CSD of Component 1



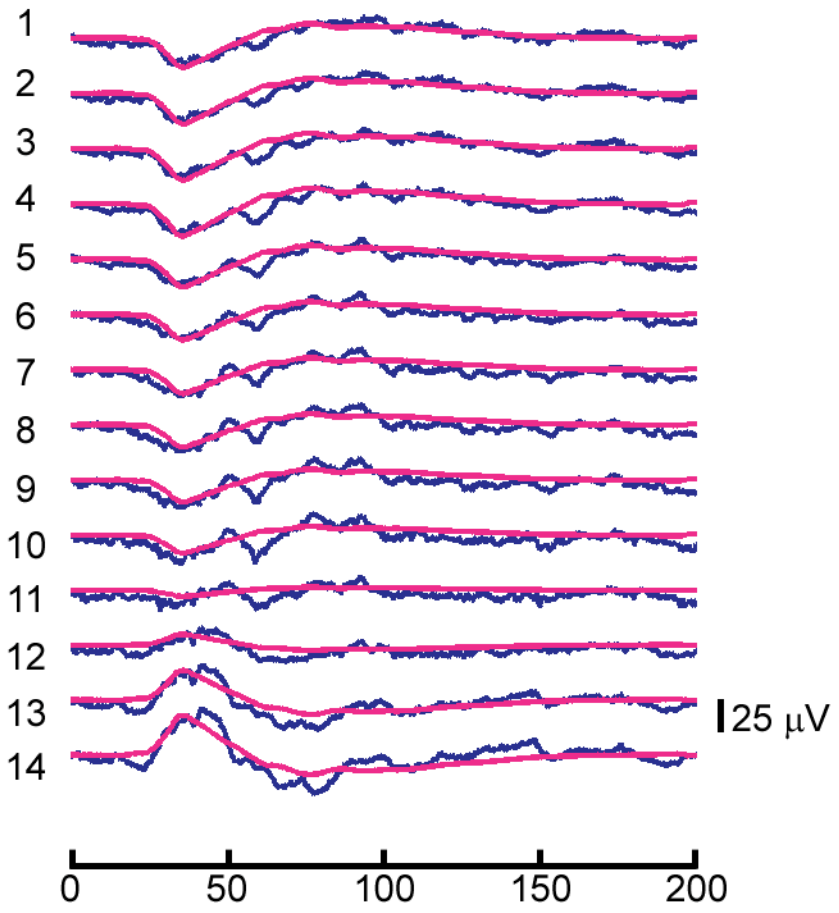
Not surprisingly, the single component model captures much of the character of the average.

How Reasonable is the Model?

CSD of Component 1

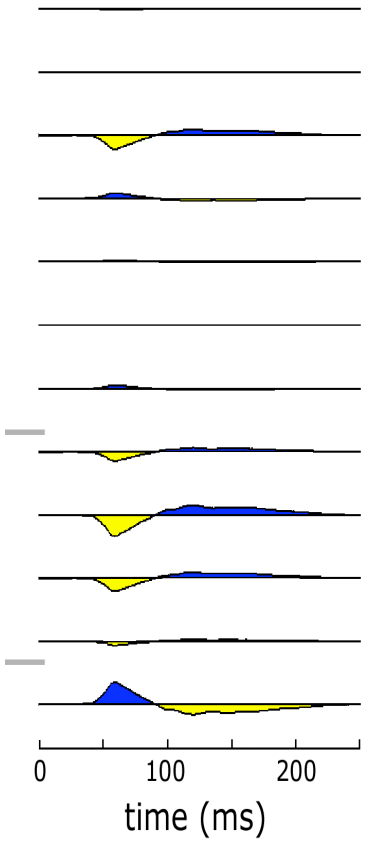


Trial 10

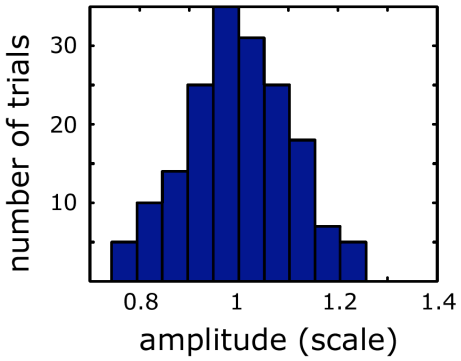


Single-Trial Characteristics

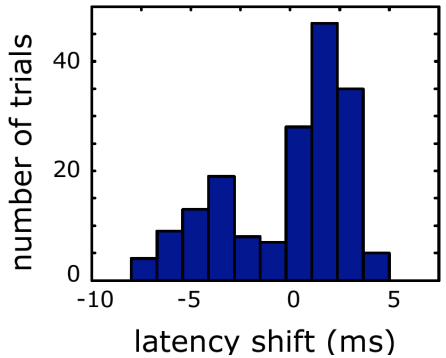
Component 1
CSD



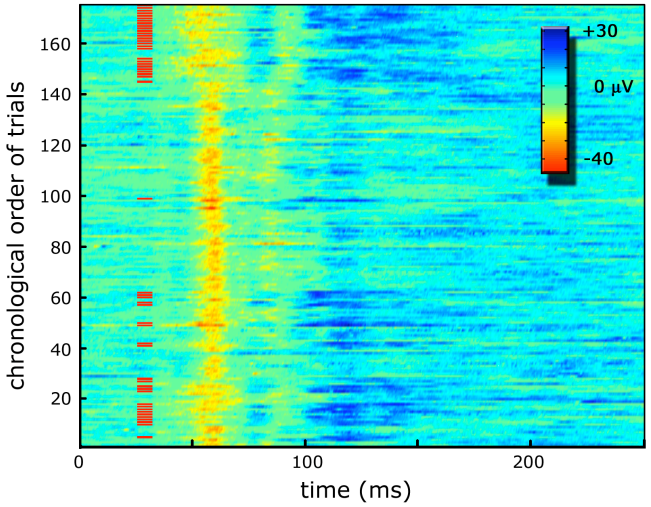
Single-Trial
Amplitudes



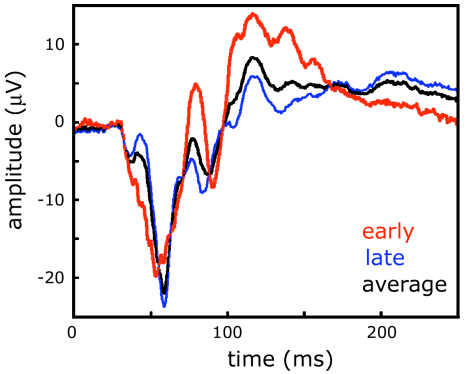
Single-Trial
Latencies



Individual Trials

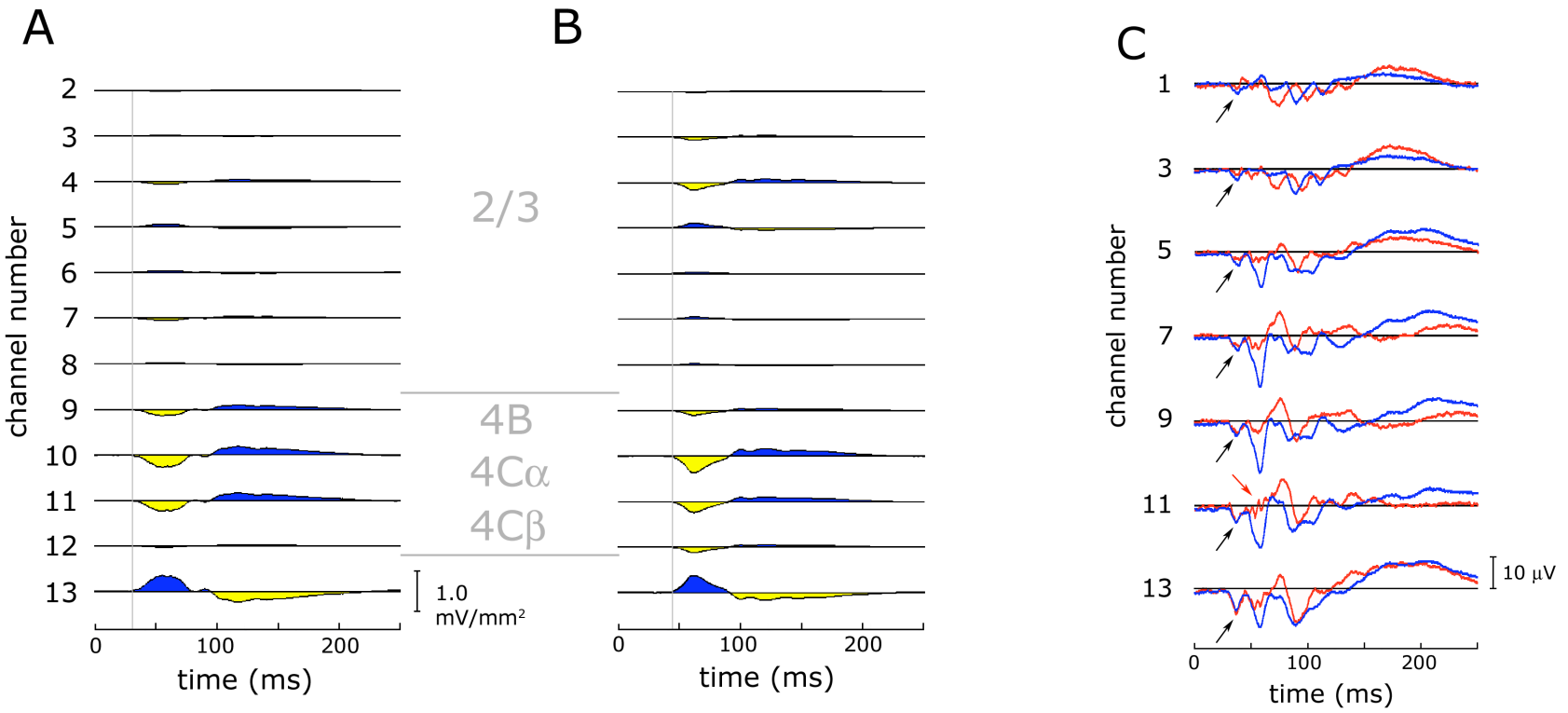


Early and Late
Subaverages



Splitting the Dataset

Splitting the dataset reveals the way in which the two response states differ

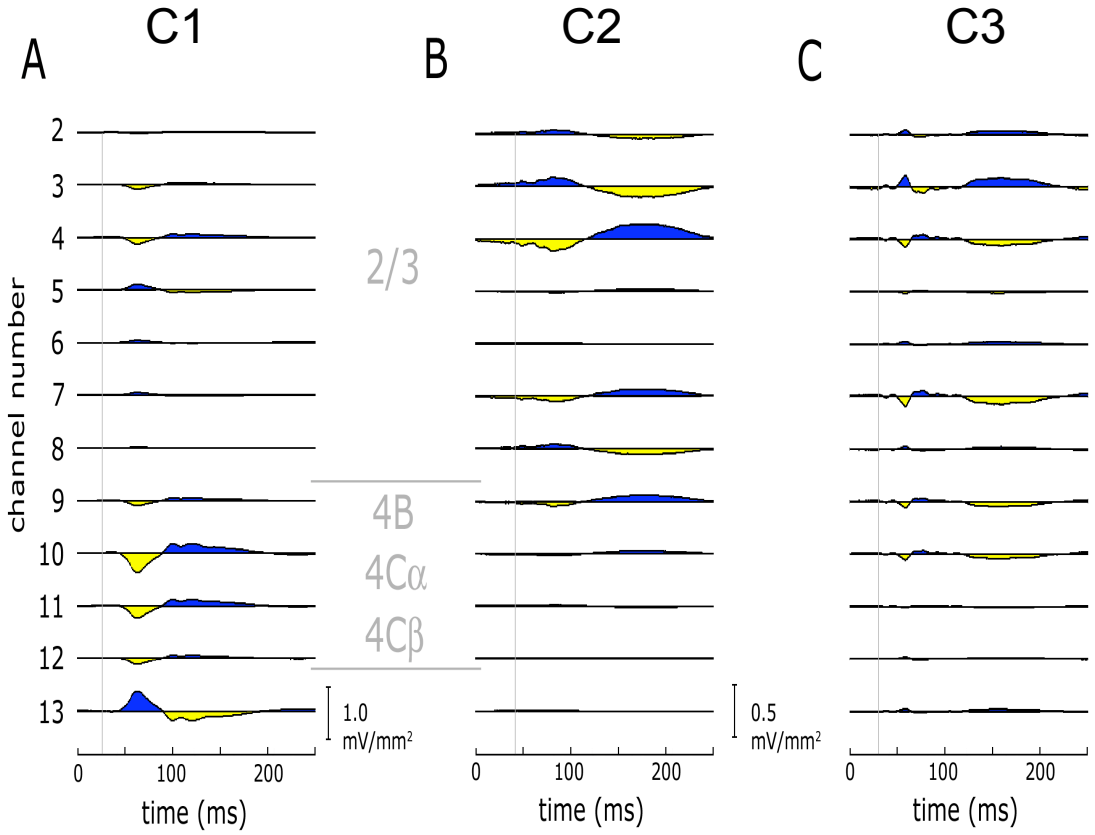


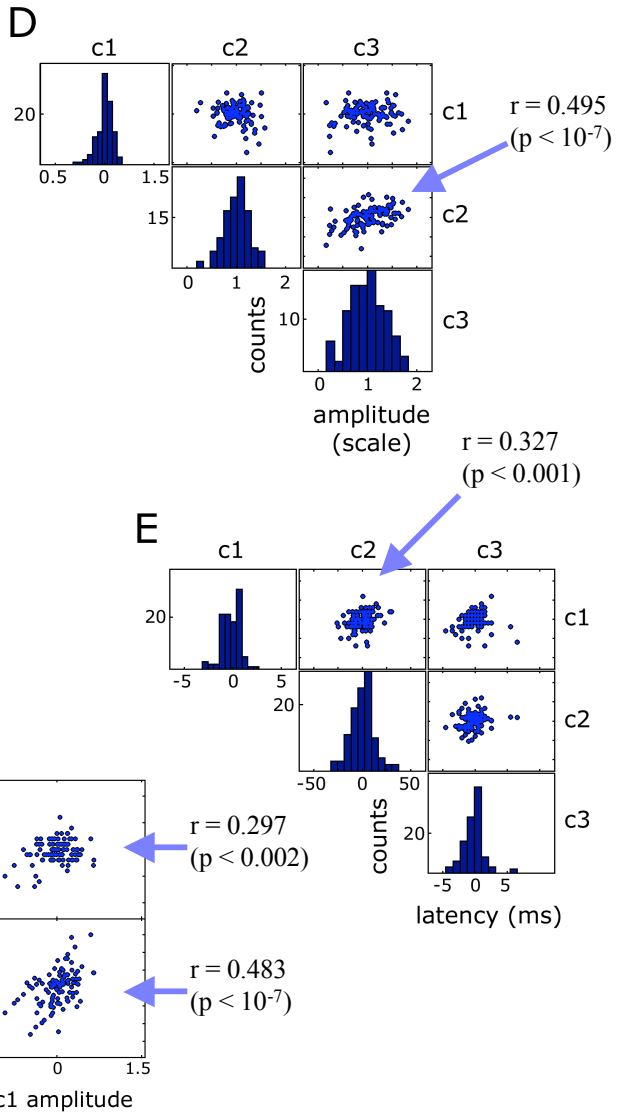
Early Waveshape

Late Waveshape

Residuals (early + late)

Three Components from the Late Subset





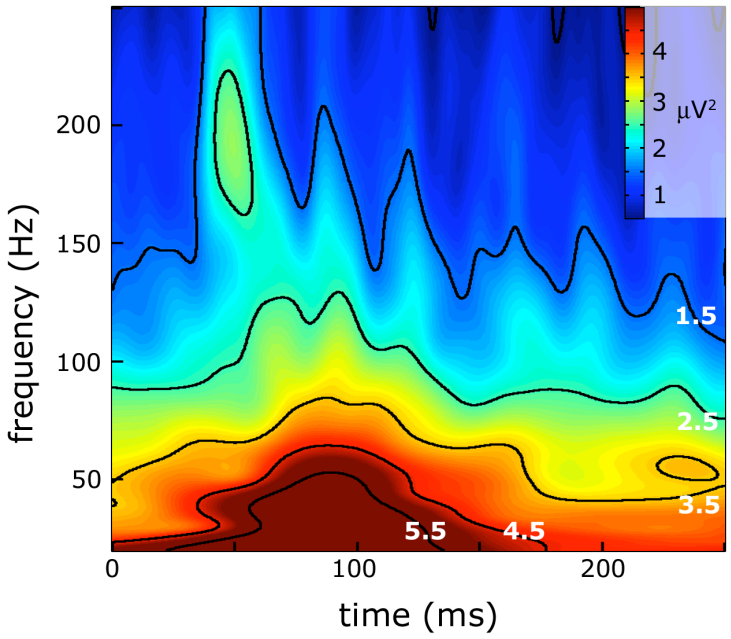
C1 and C2 are coupled in latency
C2 and C3 are coupled in amplitude
C1 amplitude and C1 latency are coupled

Studying Single-Trial Oscillations

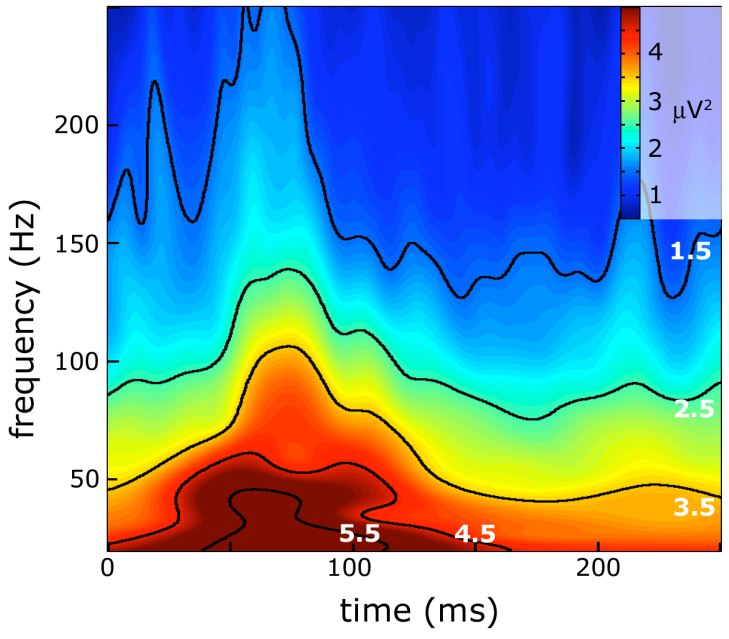
We have developed some new techniques to study oscillatory bursts, which we have found are present over a wide range of frequencies.

For example, the early and late subsets show a difference in the presence of 160-220 Hz bursts between 42-57 ms.

Early Subset



Late Subset

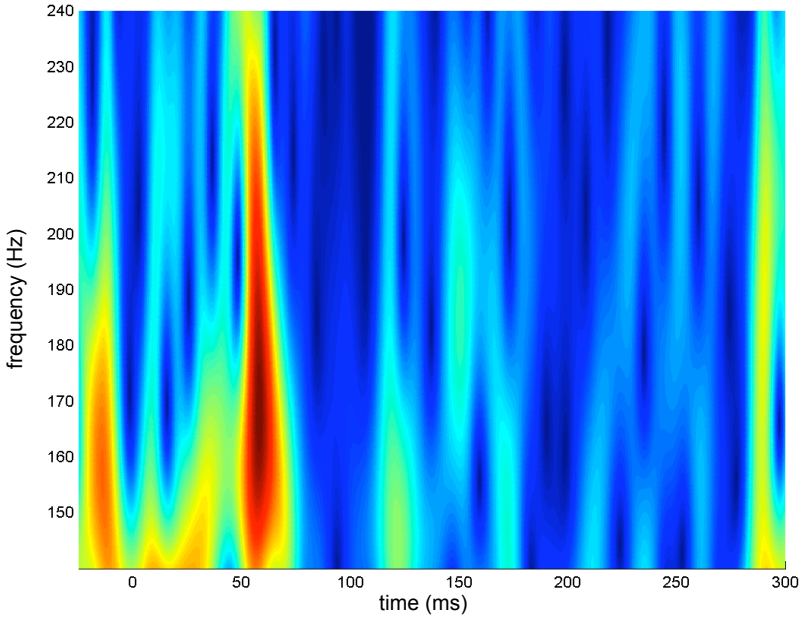


Characterizing Oscillatory Bursts

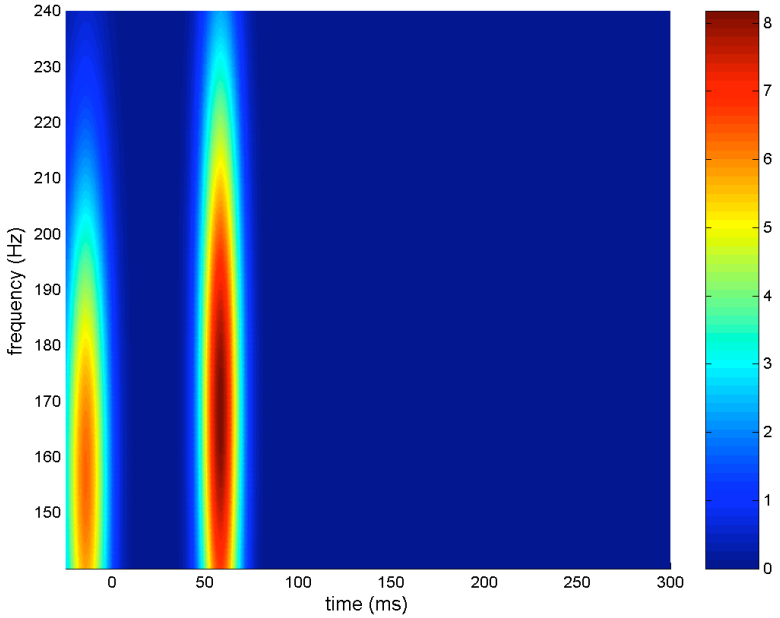
Using dVCA to identify components in the single trial allows us to remove them from the data, leaving unmodeled ongoing activity.

We can then identify oscillatory bursts and characterize them.

Time-Frequency plot of Early Subset



Modeled Bursts in the same trial



Ongoing Challenges

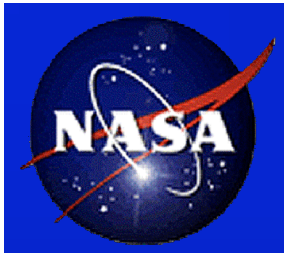
dVCA is not fast. There is a lot of computation that goes on. The more data and channels, the slower the analysis gets. High dimensional spaces are notoriously difficult to search. Advances in optimization are extremely relevant to this work.

Difficulties with large numbers of components.

The danger of getting trapped in local solutions of the model space becomes greater with increasing numbers of components. We have found that it is best to use dVCA little by little to learn about the signals.

Waveshape changes.

We do not accommodate waveshape changes. Although we have some ideas that we are currently working on to solve this.



Collaborators

Ankoor S. Shah	Albert Einstein College of Medicine, Bronx NY, USA
Wilson Truccolo	Brown University, Providence RI, USA
Charles E. Schroeder	Nathan Kline Institute, Orangeburg NY, USA
Peter Lakatos	Nathan Kline Institute, Orangeburg NY, USA
Mingzhou Ding	Florida Atlantic University, Boca Raton FL, USA
Steven L. Bressler	Florida Atlantic University, Boca Raton FL, USA

Support

NASA IDU/IS/CICT Program (KHK)
NASA Aerospace Technology Enterprise (KHK)
NARSAD Young Investigator Award (KHK)
Medical Scientist Training Program (T32M07288) (AS)
NIMH MH060358 (CES)
NSF (IBN0090717) (SLB) and (MD)
NIMH MH64204 and NIMH MH42900 (SLB) and (MD)

Thanks

We thank Drs. Ashesh Mehta and Istvan Ulbert for data collection, and Dr. Len Trejo and Sam Clanton of NASA Ames Research Center for their assistance in optimizing the coding of portions of the algorithm.

Extras

Obtaining the Marginal Posterior

Marginalizing over the variance we get the marginal posterior

$$p\left(\mathbf{C}, \mathbf{s}(t), \mathbf{\acute{a}}, \mathbf{\hat{o}} \mid \mathbf{x}(t), I\right) \propto Q^{\frac{-MRT}{2}} p(\mathbf{C} \mid I)\, p(\mathbf{s} \mid I)$$

Like before we look at the Log probability

$$\log P = -\frac{MRT}{2} \log Q + const$$

From here we look at the derivative of the Log probability wrt each of our model parameters. Setting the derivative equal to zero gives us the most probable solution.

Estimating the Waveshapes

For the j^{th} component at time q, we have

$$\frac{\partial \log P}{\partial s_j(q)} = -\frac{MRT}{2} Q^{-1} \frac{\partial Q}{\partial s_j(q)} \equiv 0$$

with

$$\frac{\partial Q}{\partial s_j(q)} = -2 \sum_{m=1}^{M} \sum_{r=1}^{R} \left[W C_{mj} \alpha_{jr} - \left(C_{mj} \alpha_{jr} \right)^2 s_j(q) \right]$$

where

$$W = x_{mr}(q + \tau_{jr}) - \sum_{\substack{n=1 \\ n \neq j}}^{N} C_{mn} \alpha_{nr} s_n(q - \tau_{nr} + \tau_{jr})$$

$$\hat{s}_j(q) = \frac{\sum_{m=1}^{M} \sum_{r=1}^{R} W C_{mj} \alpha_{jr}}{\sum_{m=1}^{M} \sum_{r=1}^{R} \left(C_{mj} \alpha_{jr} \right)^2}$$

Estimating the Amplitudes

For the j^{th} component during the p^{th} trial

$$\hat{\alpha}_{jp} = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}[U\,V]}{\sum_{m=1}^{M}\sum_{t=1}^{T}V^2}$$

where

$$U = \left(x_{mp}(t) - \sum_{\substack{n=1 \\ n\neq j}}^{N} C_{mn}\,\alpha_{np}\,s_n(t-\tau_{np}) \right)$$

$$V = C_{mj}\,s_j(t-\tau_{jp})$$

Estimating the Source-Channel Couplings

For the coupling between the j^{th} component and the i^{th} channel

$$\hat{C}_{ij} = \frac{\sum_{r=1}^{S}\sum_{t=1}^{T}[X\,Y]}{\sum_{r=1}^{S}\sum_{t=1}^{T}Y^2}$$

where

$$X = \left(x_{ir}(t) - \sum_{\substack{n=1 \\ n\neq j}}^{N} C_{in}\,\alpha_{nr}\,s_n(t-\tau_{nr}) \right)$$

$$Y = \alpha_{jr}\,s_j(t-\tau_{jr})$$

Estimating the Latency

For the latency of the j^{th} component during the p^{th} trial we **maximize**

$$Z = \sum_{m=1}^{M} \sum_{t=1}^{T} \left[C_{mj}\, \alpha_{jp}\, s_j(t - \tau_{jp}) \left(x_{mp}(t) - \sum_{\substack{n=1 \\ n \neq j}}^{N} C_{mn}\, \alpha_{np}\, s_n(t - \tau_{np}) \right) \right]$$

$$\hat{\tau}_{jp} = \arg\max Z$$

The Iterative Algorithm

$$\hat{C}_{ij} = \frac{\sum_{r=1}^{S}\sum_{t=1}^{T}[X\,Y]}{\sum_{r=1}^{S}\sum_{t=1}^{T}Y^2}$$

$$\hat{\tau}_{jp} = \arg\max Z$$

$$\hat{\alpha}_{jp} = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}[U\,V]}{\sum_{m=1}^{M}\sum_{t=1}^{T}V^2}$$

$$\hat{s}_j(q) = \frac{\sum_{m=1}^{M}\sum_{r=1}^{R}W\,C_{mj}\,\alpha_{jr}}{\sum_{m=1}^{M}\sum_{r=1}^{R}\left(C_{mj}\,\alpha_{jr}\right)^2}$$

0. Begin with initial guesses for the waveshapes $\hat{s}(q)$

   $\alpha_r = 1$

   $\tau_r = 0$

1. Estimate the source-detector couplings $\hat{C}_{ij}$

2. Estimate the single trial latencies τ_p

3. Estimate the single trial amplitudes α_p

4. Estimate the waveshapes $\hat{s}(q)$

5. Go to step 1.